



# THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

### The effects of recombination rate on the distribution and abundance of transposable elements

**Citation for published version:**

Dolgin, ES & Charlesworth, B 2008, 'The effects of recombination rate on the distribution and abundance of transposable elements' *Genetics*, vol 178, no. 4, pp. 2169-2177., 10.1534/genetics.107.082743

**Digital Object Identifier (DOI):**

[10.1534/genetics.107.082743](https://doi.org/10.1534/genetics.107.082743)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Publisher final version (usually the publisher pdf)

**Published In:**

*Genetics*

**Publisher Rights Statement:**

Freely available via Pub Med.

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# The Effects of Recombination Rate on the Distribution and Abundance of Transposable Elements

Elie S. Dolgin<sup>1</sup> and Brian Charlesworth

*Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3JT, United Kingdom*

Manuscript received October 3, 2007

Accepted for publication February 8, 2008

## ABSTRACT

Transposable elements (TEs) often accumulate in regions of the genome with suppressed recombination. But it is unclear whether this pattern reflects a reduction in the efficacy of selection against deleterious insertions or a relaxation of ectopic recombination. Discriminating between these two hypotheses has been difficult, because no formal model has investigated the effects of recombination under the deleterious insertion model. Here we take a simulation-based approach to analyze this scenario and determine the conditions under which element accumulation is expected in low recombination regions. We show that TEs become fixed as a result of Hill–Robertson effects in the form of Muller's ratchet, but only in regions of extremely low recombination when excision is effectively absent and synergism between elements is weak. These results have important implications for differentiating between the leading models of how selection acts on TEs and should help to interpret emerging population genetic and genomic data.

UNDERSTANDING how the proliferation of transposable elements (TEs) is kept in check remains an important theoretical and empirical problem in evolutionary genetics. TEs are widely considered to be intragenomic parasites, with considerable evidence that they are generally harmful (CHARLESWORTH *et al.* 1994; LYNCH 2007, Chap. 7). But despite their adverse effects, TEs represent a significant proportion of the nuclear DNA of many organisms and contribute greatly to genome evolution (KIDWELL 2002). The selfish DNA theory posits that the abundance of TEs is regulated by a balance between the tendency for elements to increase in number by transposition and natural selection against the deleterious effects they impose upon their hosts (DOOLITTLE and SAPIENZA 1980; ORGEL and CRICK 1980). However, the nature of the selective forces that limit TE abundance is still poorly understood.

The deleterious effects of TEs have been attributed to three main sources: mutations resulting from insertions into genes or regulatory sequences (“deleterious insertion model”; FINNEGAN 1992); chromosomal rearrangements caused by ectopic recombination between elements in nonhomologous insertion sites (“ectopic exchange model”; MONTGOMERY *et al.* 1987); and direct costs due to transposition activity itself (“deleterious transposition model”; BROOKFIELD 1991). An expected consequence of selection acting to limit element proliferation is that TEs should be more abundant in

regions of the genome where they are less likely to be deleterious and/or where natural selection is less effective at removing them. However, part of the problem in discriminating between the different models is that all three make similar and nonmutually exclusive qualitative predictions regarding the genomic distribution of elements. Despite extensive theoretical and experimental results, no general consensus exists on the relative importance of these different factors (BIÉMONT *et al.* 1997; CHARLESWORTH *et al.* 1997; NUZHIDIN 1999).

Many studies have tried to test the models by comparing the distribution and abundance of TEs in low *vs.* high recombination regions of the *Drosophila* genome. In the genomes of *Drosophila melanogaster* and *D. buzzatii*, high TE densities have been found in areas of reduced recombination (BARTOLOMÉ *et al.* 2002; RIZZON *et al.* 2002; BERGMAN *et al.* 2006; CASALS *et al.* 2006; FONTANILLAS *et al.* 2007). These genomic data are consistent with *in situ* hybridization data in *Drosophila* that show TEs accumulating in regions where recombination is suppressed, such as the proximal and distal arms of the major autosomes (LANGLEY *et al.* 1988; CHARLESWORTH *et al.* 1992; MASIDE *et al.* 2001; but see HOOGLAND and BIÉMONT 1996), in polymorphic chromosomal inversions (EANES *et al.* 1992; SNIEGOWSKI and CHARLESWORTH 1994), and on the neo-Y chromosome of *D. miranda* (STEINEMANN and STEINEMANN 1998; BACHTROG 2003). Since areas of low meiotic recombination are expected to have low rates of ectopic recombination (LANGLEY *et al.* 1988; MONTGOMERY *et al.* 1991; GOLDMAN and LICHTEN 1996, 2000), these results may be taken as evidence for the ectopic exchange model.

<sup>1</sup>Corresponding author: Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, King's Bldgs., Edinburgh EH9 3JT, UK. E-mail: elie.dolgin@ed.ac.uk

However, these results are also consistent with the deleterious insertion model, since low recombination regions have lower gene densities (ADAMS *et al.* 2000; FULLERTON *et al.* 2001) and experience weaker effective selection against deleterious insertions (CHARLESWORTH and LANGLEY 1991). Similarly, TE accumulation in low recombination regions is expected under the deleterious transposition model, which makes comparable predictions regarding element patterning to the deleterious insertion model (BROOKFIELD 1991). Therefore, all three models predict a greater abundance of elements in genomic regions of reduced recombination, and no firm conclusions can be drawn from this observation alone as to whether the accumulation of TEs in low recombination regions reflects differences in the rate of ectopic exchange or a reduction in the efficacy of selection acting against deleterious element insertions.

While a strong theoretical framework exists for predicting the effects of recombination rate on ectopic exchange and its implications for TE distribution and abundance (LANGLEY *et al.* 1988; CHARLESWORTH *et al.* 1992), no theoretical studies have taken into account the effects of close linkage on the efficacy of selection under the deleterious insertion model. Local recombination should be an important factor affecting element patterning because a deleterious TE at one site can reduce the efficacy of selection acting at neighboring linked insertion sites, a phenomenon known as the Hill–Robertson effect (HILL and ROBERTSON 1966; FELSENSTEIN 1974). This reduces the effective population size, enhances the importance of genetic drift, and can lead to the fixation of TEs in nonrecombining regions through the process of Muller’s ratchet (MULLER 1964; FELSENSTEIN 1974). These effects should be influenced by a number of factors, including local recombination rates, effective population sizes, the strength of selection, and the rate of occurrence of new insertions. These factors have been studied for classical mutational models (*e.g.*, GORDO and CHARLESWORTH 2000, 2001; McVEAN and CHARLESWORTH 2000; COMERON and KREITMAN 2002; KEIGHTLEY and OTTO 2006), but the unique ability of transposable elements to replicate autonomously through transposition, and excise from the genome, could lead to distinct patterns and evolutionary dynamics. A better understanding of the genomic distribution expected from these effects should help differentiate the deleterious insertion model from the ectopic exchange model.

Theoretical considerations of TEs under the deleterious insertion model have almost universally assumed a constant recombination rate across the entire genome. Part of the reason that the effects of variable recombination have not been well-studied is that understanding these processes often requires a simulation-based approach. This is in contrast to the relatively simple prediction of the ectopic exchange model that the number of TE copies inserted in a given region is

expected to be inversely related to the recombination rate (LANGLEY *et al.* 1988; CHARLESWORTH *et al.* 1992). In this article, we develop a simulation model that simultaneously considers a genome with both high and low recombination regions. We then identify conditions and parameters under which element accumulation in regions of reduced recombination is expected. We show that the deleterious insertion model can explain an increased abundance of TEs in low recombination areas only when individual element insertions become fixed at the population level. We discuss the results of our simulations in relation to the genomic data and population surveys of *D. melanogaster* to try and distinguish between the leading hypotheses for why TEs build up in low recombination regions.

## METHODS

Computer simulations were used to examine the effects of reduced recombination on the distribution and abundance of transposable elements in the genome. The model was implemented by modifying the simulation program of DOLGIN and CHARLESWORTH (2006) to incorporate sexual reproduction and recombination. C++ files are available upon request. Random numbers were generated using the Mersenne Twister pseudorandom number generator (MATSUMOTO and NISHIMURA 1998), adapted for C++ by J. Bedaux (<http://www.bedaux.net/mtrand/>).

**Population and genome:** We consider a diploid population of a constant size,  $N$ , with discrete non-overlapping generations. The genome is composed of four chromosomes: three large recombining chromosomes, each representing 30% of the genome, and one smaller chromosome, representing 10% of the genome. This is intended to provide a rough portrayal of the *D. melanogaster* genome, for which most data on TE distributions are available, although we have exaggerated the size of the small, noncrossing over fourth chromosome (ASHBURNER *et al.* 2004). The larger chromosomes are able to carry as many as 1000 elements each, permitting a large number of potential insertions, with a recombination frequency,  $r$ , of  $10^{-3}$  between any two adjacent potential insertion sites, resulting in one crossover event per chromosome per generation on average. The smaller chromosome can carry a maximum of 333 elements, and the recombination frequency is adjusted in different simulations to analyze the effects of a genomic region with reduced and null recombination, respectively.

**Simulations:** The simulation is initiated by randomly inserting  $N$  elements into the population (*i.e.*, 1 copy per genome on average). Each generation then starts with reproduction, involving selection and recombination, followed by transposition and excision. In each generation, individuals are sampled as follows to pro-

duce a new offspring population of size  $N$ . For each offspring, two parents are randomly selected. Each parent then creates a gamete, with the number of crossovers per chromosome drawn from a Poisson distribution according to the length and recombination rate of the chromosome and a uniform distribution of crossover positions. The gametes are then combined, and the resulting offspring are subject to selection, with the probability of survival proportional to their fitness values, as described by DOLGIN and CHARLESWORTH (2006). Fitness is assumed to be a decreasing function of TE abundance. This is essential to allow for a stable equilibrium copy number under free recombination; otherwise, TEs will proliferate in an unbounded fashion regardless of recombination rate (CHARLESWORTH and CHARLESWORTH 1983). We model the fitness of an individual with  $n$  elements by an exponential quadratic, decreasing function of the TE copy number,

$$w_n = \exp(-an - \frac{1}{2}bn^2), \quad (1)$$

where  $a$  and  $b$  are constant selection coefficients (CHARLESWORTH 1990), and dominance is assumed to be intermediate at each insertion site (*i.e.*,  $n$  is the sum of all the elements in the diploid genome). The model permits TEs to segregate in the population and implicitly assumes that insertions causing large fitness effects are rapidly eliminated by selection and can be ignored (CHARLESWORTH and LANGLEY 1991). After formation of the new offspring population, transposition and excision are assumed to occur randomly throughout each genome under a Poisson process, with the TE copy number increasing with probability  $u$  per element and decreasing with probability  $v$  per element.

**Model parameters:** Simulations were generally run for  $10^6$  generations, although in cases where nonequilibrium TE accumulation was especially rapid, simulations were run for fewer generations. At every 1000th generation, we surveyed the TE copy number on each chromosome and the number of sites where a TE was fixed across the entire population. At least 10 simulations were performed for each combination of parameters investigated. We set the parameters in our model to match estimates from *Drosophila* populations, where average transposition rates for all classes of elements are  $\sim 10^{-4}$  per element per generation, with excision rates at least one order of magnitude smaller (NUZHIDIN and MACKAY 1995; SUH *et al.* 1995; VIEIRA and BIÉMONT 1997; PASYUKOVA *et al.* 1998; MASIDE *et al.* 2000). Since equilibrium requires equality of rates of origination and removal of elements, the strength of selection against segregating elements should be of the order  $10^{-5}$ – $10^{-4}$  per copy (CHARLESWORTH *et al.* 1994). Because of synergism between elements, the strength of selection depends on both  $a$  and  $b$ , so we set  $a = 10^{-5}$  and let  $b$  vary from  $10^{-6}$  to  $10^{-5}$ .

However, to achieve more simulations in a given amount of computation time, we adjusted the parameters following the scaling employed and validated in the simulations of DOLGIN and CHARLESWORTH (2006), which has also been used elsewhere (LE ROUZIC *et al.* 2007; SÖDERBERG and BERG 2007). This scaling involves increasing the transposition–excision and selection parameters ( $u$ ,  $v$ ,  $a$ , and  $b$ ) by two orders of magnitude, while decreasing  $N$  and the number of generations by the same factor. Because we are maintaining the products of the deterministic forces and the effective population size constant, this should lead to the same evolutionary outcome, provided that all evolutionary forces are weak (EWENS 2004, Chap. 4). In the results below, we present the uncorrected parameter values as used in the simulations, although to reflect biologically relevant values, all the parameters would need to be scaled 100-fold.

## RESULTS

With the quadratic fitness formula (CHARLESWORTH 1990; DOLGIN and CHARLESWORTH 2006), the expected equilibrium mean TE copy number under free recombination with low frequencies of elements at each occupied site is given by

$$\bar{n} \approx \frac{u - a - v}{b}. \quad (2)$$

With the standard rate of crossing over across our entire simulated genome, including the small (“fourth”) chromosome, we found that the TE copy number in simulations of large populations ( $N = 10^3$ ) equilibrated rapidly and matched this predicted amount within  $\pm 1$  element for all parameter combinations tested (results not shown). In the absence of excision, when we set the recombination rate on the fourth chromosome equal to zero, equilibrium copy numbers as predicted by Equation 2 were still established when the synergism coefficient,  $b$ , was large, implying strong synergistic interactions between insertions. Under these conditions,  $\sim 10\%$  of the total number of elements in the genome were on the nonrecombining fourth chromosome, which is directly proportional to the size of the chromosome in our simulations. However, with lower values of  $b$ , the genome accumulated more elements than the equilibrium expectations (Figure 1). This is the result of a buildup of elements on the nonrecombining fourth chromosome, as TEs become fixed due to Muller’s ratchet (MULLER 1964; FELSENSTEIN 1974). The shift between TEs either equilibrating or accumulating on the nonrecombining chromosome occurs quite suddenly following slight changes in the synergism coefficient,  $b$ . By analyzing a single simulation run, we can observe this shift and see these fixations as discrete jumps in the TE copy number on the fourth chromosome between periods when the mean TE copy number remains fairly constant (Figure 1). This is in contrast to

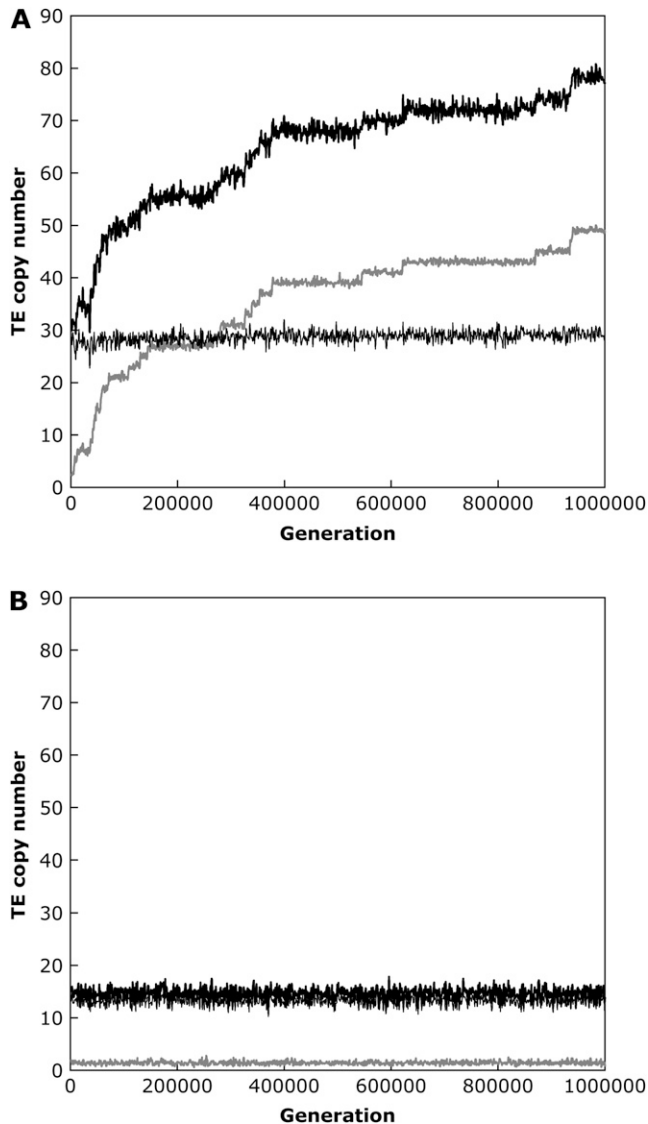


FIGURE 1.—An example simulation run with weak (A) and strong (B) synergism when the fourth chromosome is completely nonrecombining. Diploid mean TE copy numbers are shown for the entire genome (thick black line), the nonrecombining fourth chromosome (thick gray line), and the three freely recombining chromosomes (thin black line) for simulation parameters  $N = 1000$ ,  $u = 0.01$ ,  $v = 0$ ,  $a = 0.001$ , and  $b = 0.0003$  (A) or  $b = 0.0006$  (B).

the freely recombining chromosomes, which maintained stable equilibrium copy numbers throughout the simulations, and the situation with strong synergism between elements, where no fixations were observed (Figure 1).

Classical mutation models usually assume a uniform genomic mutation rate and a constant selection coefficient for new mutations, resulting in a constant expected speed of Muller's ratchet (*e.g.*, GORDO and CHARLESWORTH 2000, 2001; STEPHAN and KIM 2002). On the other hand, population genetic models of transposable elements generally assume a fixed rate of transposition per element (LE ROUZIC and DECELIERE 2005), leading to

greater genomic transposition rates with higher TE copy numbers. To counter this potential proliferation, the fitness effects of each insertion must be a decreasing function of TE abundance (CHARLESWORTH and CHARLESWORTH 1983). Thus, with more elements, two antagonistic forces can act: the ratchet may be either accelerated as a result of the increased genomic transposition rate or decelerated owing to stronger selection against insertions.

Under the quadratic fitness equation used here, we found that the time between successive fixations generally increased as copy number increased, indicating that Muller's ratchet slows down as more elements accumulate on the nonrecombining chromosome, and suggesting that the effects of stronger selection outweigh the higher genomic transposition rate (Figure 2). For example, for the two sets of parameters investigated in Figure 2, we performed a least-squares linear regression of time between fixations and "click" number, and found significantly positive slopes (Figure 2A,  $F_{1,465} = 56.3$ ,  $P < 0.0001$ ; Figure 2B,  $F_{1,209} = 7.42$ ,  $P = 0.007$ ). Our analysis is different from GORDO and CHARLESWORTH's (2001) model of Muller's ratchet and transposable elements, as they essentially modified a mutational model to fit estimates of constant genomic transposition rates and assumed multiplicative fitness effects. However, it is consistent with an asexual mutational model that found that Muller's ratchet can effectively be halted if deleterious mutations have synergistic fitness effects (KONDRASHOV 1994; but see BUTCHER 1995), although in our model we never observed the fourth chromosome to reach a stable equilibrium copy number in simulations where TE fixations occurred, even over millions of generations (which equates to hundreds of millions of generations after scaling the parameters appropriately).

If we ignore the fact that the rate of Muller's ratchet is not constant, for the sake of comparison, and simply calculate a mean number of fixations per generation across the entire length of the simulation run, we have a rough measure with which to compare the speed with which TEs accumulate under different parameter conditions. Table 1 shows this average rate of fixation on the fourth chromosome under a range of synergism coefficients for two population sizes. As in Figure 1, it can be seen that element fixation was much more prevalent when synergism was weak (*i.e.*,  $b$  was small). For a given strength of synergism, element fixation was also more likely to occur when populations were smaller, and fixations tended to be more rapid in smaller populations. We also considered the situation when local recombination was reduced but not completely absent. When the recombination frequency on the fourth chromosome was 10-fold lower than the rest of the genome ( $r = 10^{-4}$  between any two insertion sites, resulting in one crossover event per fourth chromosome every  $\sim 30$  generations on average), no fixations were ever observed for any set of parameters tested (results not shown). With a

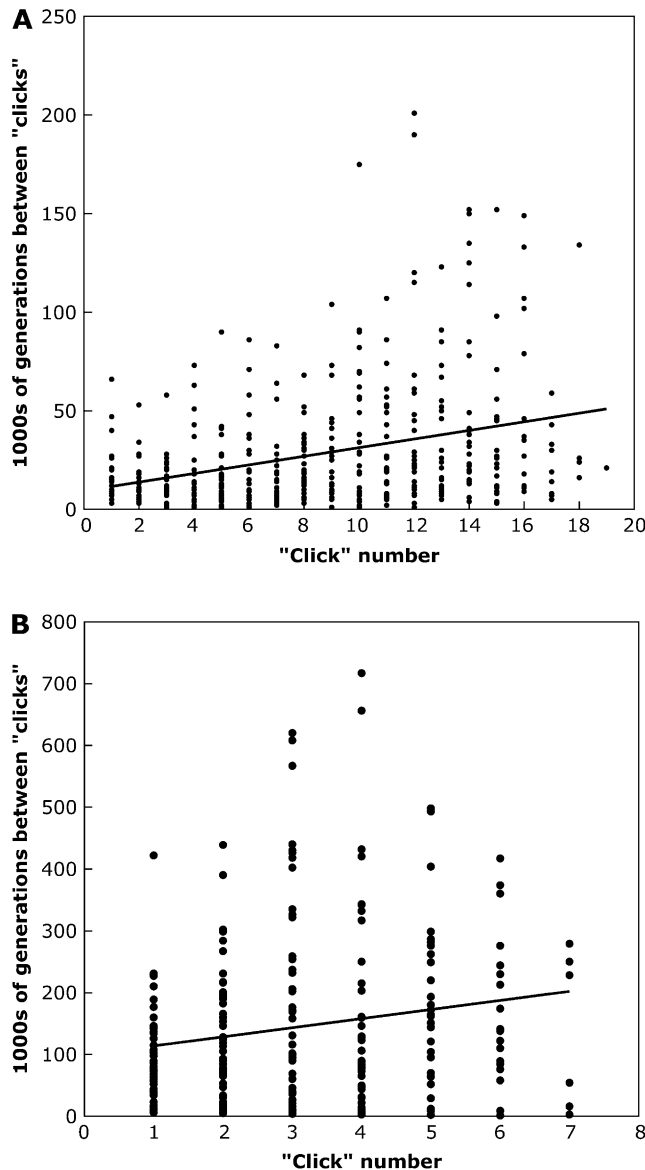


FIGURE 2.—The time between clicks of the ratchet as a function of the ratchet number for simulations with a completely nonrecombining fourth chromosome for simulation parameters:  $N = 1000$ ,  $u = 0.01$ ,  $v = 0$ ,  $a = 0.001$ , and  $b = 0.0003$  (A) or  $b = 0.0004$  (B). A shows the results of 30 simulations run for  $5 \times 10^5$  generations, and B shows the results of 50 simulations run for  $10^6$  generations. The line shows the least-squares linear regression.

100-fold lower recombination frequency on the fourth chromosome ( $r = 10^{-5}$  between any two insertion sites, resulting in one crossover event per fourth chromosome every  $\sim 300$  generations on average), element accumulation was widespread, but occurred at a considerably slower rate than in the complete absence of recombination (Table 1).

Our model thus far has assumed no element excision; therefore, once an element becomes fixed, Muller's ratchet has clicked irreversibly. While this lack of excision may be true for some classes of retrotransposons, which

show little or no excision, the cut-and-paste transposition mechanism of DNA-based elements often leads to frequent element loss (CRAIG *et al.* 2002; HUA-VAN *et al.* 2005). We incorporated excision into our model, and inspected the effects of excision rates one to three orders of magnitude below the rate of transposition for a genome with a completely nonrecombining fourth chromosome. For most parameter combinations tested, the presence of element excision halted the process of Muller's ratchet (Figure 3). Although temporary fixations were observed, these elements eventually became excised, allowing the population to reach a stable equilibrium close to that predicted by Equation 2. Only when the rate of fixation was much greater than the rate of excision did Muller's ratchet outpace excision, leading to continual element accumulation on the fourth chromosome (*i.e.*, with  $b = 0.0002$  and  $v = 10^{-5}$ , see Figure 3).

## DISCUSSION

Many authors have suggested that the reduction in the pressure of selection against TEs when recombination is reduced can lead to the buildup of elements in low recombination regions (*e.g.*, CHARLESWORTH and LANGLEY 1991; DURET *et al.* 2000; BAROLOMÉ *et al.* 2002; RIZZON *et al.* 2002; FONTANILLAS *et al.* 2007); here we provide the first theoretical analysis of this situation. We have shown that accumulation of TEs under the deleterious insertion model is expected to occur only in the genomes of small populations where recombination is severely reduced, when excision is extremely rare or absent, and with weak synergism between elements. These criteria should be empirically testable, yet might be satisfied only under a very limited range of biological circumstances. This is in contrast to the ectopic exchange model, which broadly predicts that the number of TE insertions is inversely related to the recombination rate, even in very large populations (LANGLEY *et al.* 1988; MONTGOMERY *et al.* 1991; CHARLESWORTH *et al.* 1992). Therefore, if TE accumulation is observed in low recombination regions, but the conditions necessary for buildup under the deleterious insertion model are not met, we can infer that the ectopic exchange model is implicated. Our model thus provides a first step in distinguishing between the leading models of TE distribution and abundance in regions of different recombination.

The effects of recombination on the deleterious insertion model were previously investigated only without any variation in recombination rate across the genome. CHARLESWORTH and CHARLESWORTH (1983) identified only minor effects of the recombination rate on TE abundance, and DOLGIN and CHARLESWORTH (2006) found large effects of the complete absence of recombination with asexuality, but no theoretical study has considered chromosomal regions with different recombination rates. This is important, because differ-

TABLE 1  
The speed of Muller’s ratchet on a null- or low-recombining chromosome

<i>b</i>	Null recombination				Low recombination			
	<i>N</i> = 1000		<i>N</i> = 500		<i>N</i> = 1000		<i>N</i> = 500	
	Clicks	Runs	Clicks	Runs	Clicks	Runs	Clicks	Runs
0.0002	39.4	20 <sup>a</sup>	—	—	12.9	10 <sup>b</sup>	—	—
0.0003	3.1	30 <sup>c</sup>	—	—	0.50	10	9.95	10 <sup>b</sup>
0.0004	0.49	50	7.4	20 <sup>b</sup>	0.03	10	1.34	10
0.0005	0.12	30	1.86	10 <sup>c</sup>	0	10	0.55	10
0.0006	0.01	30	0.62	10			0.22	10
0.0007	0	10	0.34	10			0.13	10
0.0008			0.15	10			0.02	10
0.0009			0.17	10				

The average number of clicks per 10<sup>5</sup> generations is shown, assuming a constant rate of element fixation, with the recombination rate on the fourth chromosome either 0 (null recombination) or 10<sup>−5</sup> between insertion sites (low recombination). The remaining parameters are *u* = 0.01, *v* = 0, *a* = 0.001.

All simulations are run for 10<sup>6</sup> generations, except where indicated. — denotes simulations where element fixation was so rapid that the computing power was insufficient to run the simulation for 10<sup>5</sup> generations.

<sup>a</sup> 10<sup>5</sup> generations.  
<sup>b</sup> 2 × 10<sup>5</sup> generations.  
<sup>c</sup> 5 × 10<sup>5</sup> generations.

ences in the local recombination rate can drastically influence the effective population size experienced by insertions, as a result of the Hill–Robertson effects (HILL and ROBERTSON 1966; FELSENSTEIN 1974). The importance of effective population size for the abundance of TEs in general was first noted by CHARLESWORTH and CHARLESWORTH (1983), who found that smaller populations led to slightly greater equilibrium TE copy numbers in the presence of recombination. Later studies showed that small population size with sexual reproduction can often lead to unlimited increases in copy number, especially with high selfing rates, but the population sizes involved in such situations were unrealistically small for most natural populations (BROOKFIELD and BADGE 1997; WRIGHT and SCHOEN 1999).

Here we have used large, randomly mating populations that may be more representative of natural populations of organisms like *Drosophila*, focusing on TEs with weak fitness effects that can segregate in noncoding regions. We found that Hill–Robertson effects in low recombination regions can lead to stochastic TE accumulation. Even so, the population sizes used in this study are still probably smaller than those found in nature for *Drosophila* populations. Most of our simulations involved a population size of *N* = 10<sup>3</sup>, which reflects a size of 10<sup>5</sup> following the scaling employed here (see METHODS). This is an order of magnitude lower than the effective population size of 10<sup>6</sup> typically assumed for *D. melanogaster* (KREITMAN 1983). Simulations run with larger populations indicated that TE accumulation was less likely and/or occurred at a slower rate (results not shown), but these simulations were too slow and computationally intensive to replicate many

times. The possibility of TE accumulation in natural populations could thus be more limited than the already restrictive set of circumstances identified in our model, where larger population sizes reduce Hill–Robertson

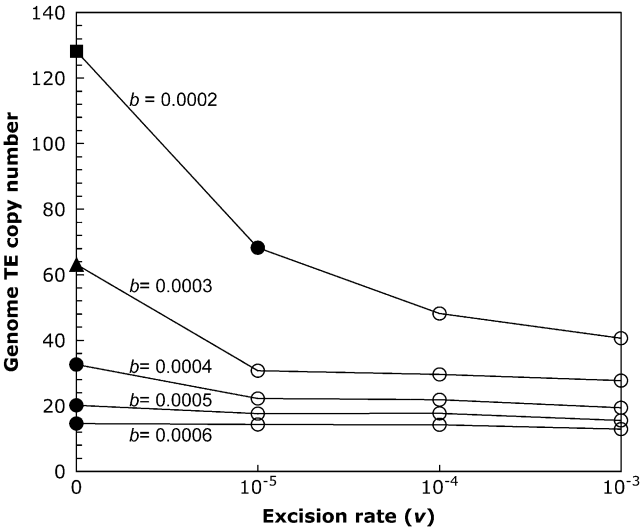


FIGURE 3.—The effect of excision on element abundance. The mean genome TE copy number at the end of the simulation is shown for simulations with a completely nonrecombining fourth chromosome for a range of synergism coefficients and excision rates, and the remaining simulation parameters were held constant: *N* = 1000, *u* = 0.01, *a* = 0.001. Simulations were run for 10<sup>6</sup> generations (denoted by circles), except for the triangle (5 × 10<sup>5</sup> generations) and the square (10<sup>5</sup> generations). Open circles indicate simulations that reached stable transposition–excision–selection equilibria, and solid symbols indicate simulations where fixations leading to nonequilibrium dynamics were observed. All data points are based on at least 10 simulation runs.

effects. On the other hand, our model did not consider other factors that might reduce the effective population size and potentially accelerate the rate of ratchet, such as population subdivision (COMBADO *et al.* 2007) or effects of background selection and selective sweeps at linked coding or regulatory sites (GORDO and CHARLESWORTH 2001). The latter effect could well be very important for nonrecombining genomic regions like the fourth chromosome (SHELDAHL *et al.* 2003) and the neo-Y chromosome of *D. miranda* (BACHTROG 2003), where DNA sequence diversity is lower than the genomewide average by more than a factor of 10, implying a large reduction in effective population size.

Another problem in relating our results to *Drosophila* data is that the nonrecombining fourth chromosome in our model represented 10% of the genome. In reality, however, the fourth chromosome is much smaller. When we adjusted the chromosome sizes and made the fourth chromosome occupy only 1% of the genome, TE accumulation was never observed under any of the parameter conditions permitting accumulation on the larger chromosome that we used (results not shown). This suggests that selection is more effective at preventing TE buildup when the insertion rate of new elements into nonrecombining regions is lower, consistent with general Muller's ratchet theory that fixations occur more rapidly with greater mutation rates (GORDO and CHARLESWORTH 2000). Although we exaggerated the size of the fourth chromosome, a larger chromosome might more accurately approximate other nonrecombining chromosomes including sex chromosomes, such as the neo-Y chromosome in *D. miranda*, which is riddled with TEs (STEINEMANN and STEINEMANN 1998; BACHTROG 2003). Our model also considered only genomic element copy numbers in the tens to hundreds, but the total number of active elements in the *Drosophila* genome is probably about an order of magnitude greater (BARTOLOMÉ *et al.* 2002; KAMINKER *et al.* 2002; BERGMAN *et al.* 2006). However, given the inflated size of the fourth chromosome in our simulations, the reduced genomic copy number counterbalances the enlarged chromosome size and makes the rate of new insertions into the fourth chromosome reasonably realistic.

The results of GORDO *et al.* (2002) show that, when Muller's ratchet operates, the effect of deleterious mutations on neutral diversity in nonrecombining genomic regions is less than that predicted by equilibrium background selection theory. We have argued that the ratchet is unlikely to be causing a buildup of TEs on the fourth chromosome of *D. melanogaster*. It follows that their effect on neutral diversity should be adequately predicted by the background selection formula if TEs are contained by selection against the deleterious effects of element insertions on fitness (CHARLESWORTH 1996). But the predicted reduction in diversity due to background selection effects from the observed number of segregating TEs on the fourth chromosome is at least a

factor of 1000-fold (CHARLESWORTH 1996; CARR *et al.* 2001), compared with an observed autosome to fourth chromosome silent diversity ratio of about 20 for *D. melanogaster* (SHELDAHL *et al.* 2003). This suggests that, as noted by CHARLESWORTH (1996), the background selection model greatly overpredicts the reduction in diversity. One possibility is that element insertions outside coding sequences are largely neutral on the fourth chromosome, due to the absence of meiotic ectopic exchange. These elements would not, therefore, exert background selection effects. This faces the difficulty, however, that retrotransposons do not show evidence for accumulation on the fourth chromosome (BARTOLOMÉ *et al.* 2002; KAMINKER *et al.* 2002; BERGMAN *et al.* 2006), as we would expect with this hypothesis. The solution to this dilemma may be that many members of this element class appear to have invaded *D. melanogaster* recently, and have not yet had time to accumulate on the fourth chromosome (SÁNCHEZ-GARCÍA *et al.* 2005; BERGMAN and BENSASSON 2007).

Nonetheless, analyses of the *D. melanogaster* genome have revealed that many families of TEs, including both DNA-based elements and retrotransposons, tend to accumulate in euchromatic and heterochromatic regions where recombination is reduced (BARTOLOMÉ *et al.* 2002; KAMINKER *et al.* 2002; RIZZON *et al.* 2002; BERGMAN *et al.* 2006; FONTANILLAS *et al.* 2007), but the nature of the mechanisms involved is obscure. The results of our model indicate that TE accumulation under the deleterious insertion model is expected only when element excision is absent or extremely rare. Because DNA transposons can excise from the genome, albeit at a low rate (NUZHDIN and MACKAY 1995; SUH *et al.* 1995; VIEIRA and BIÉMONT 1997; PASYUKOVA *et al.* 1998; MASIDE *et al.* 2000), Muller's ratchet is likely to be ineffective for this class of elements. It thus seems more likely that the ectopic exchange model explains the *Drosophila* data, because the nonrecombining fourth chromosome is enriched in DNA-based elements with the ability to excise (BARTOLOMÉ *et al.* 2002; KAMINKER *et al.* 2002; BERGMAN *et al.* 2006). This is further supported by the observation that the density of DNA transposons parallels the recombination rate along the chromosomes (RIZZON *et al.* 2002), whereas our model shows that low to moderate amounts of recombination effectively prevent elements from accumulating.

Unlike DNA-based elements, retrotransposons rarely excise from the genome (CRAIG *et al.* 2002; HUA-VAN *et al.* 2005) and tend to be fairly homogeneously distributed, except for accumulation in some regions with extremely low recombination rates (RIZZON *et al.* 2002; but see BARTOLOMÉ *et al.* 2002). Recent analyses have also demonstrated that these elements are often fixed within populations in nonrecombining portions of the *Drosophila* genome (BACHTROG 2003; BARTOLOMÉ and MASIDE 2004; MASIDE *et al.* 2005). BARTOLOMÉ and MASIDE (2004) used a PCR-based approach to determine the population



frequency of a number of TE insertions in regions of suppressed or free recombination. They found that elements in regions of high recombination segregated at low frequencies and tended to be young, whereas null recombination regions were enriched in old elements at high frequencies or fixation, including several retrotransposons. This is at least qualitatively consistent with a model of retrotransposons accumulating and becoming fixed in low recombination regions due to Muller's ratchet. However, this model would also predict accumulation of retrotransposons on the fourth chromosome; as discussed above, this is not the case. The presence of fixed TEs is nonetheless compatible with the ectopic exchange model, because neutral elements can drift to high frequencies and fixation in low recombination regions, given that these regions have small effective population sizes.

DNA-based elements and retrotransposons show many differences in their respective intragenomic dynamics and distributions, and the causes of these differences remain unclear (HUA-VAN *et al.* 2005). Here we have shown that a selective mechanism that has often been suggested to explain the prevalence of TEs in regions of suppressed recombination can probably be ruled out for DNA-based elements, because they exhibit appreciable rates of element excision. The limited population survey data for retrotransposons in *Drosophila* is consistent with our model, but screening of more insertions in multiple populations will be needed to test whether fixations are in fact more prevalent in low recombination regions. With the sequencing of more *D. melanogaster* strains, we should be able to better understand the population level processes underlying our genomic observations. Although we have identified some necessary conditions for TEs to build up through Muller's ratchet, determining conclusively which selective mechanisms are involved in causing TEs to accumulate in low recombination regions of the genome requires more experimental and theoretical analyses.

We thank the members of the Ashworth Xgrid cluster for computing time, and L. Loewe for helpful comments on the manuscript. This work was supported by postgraduate research scholarships to E.S.D. from the Natural Sciences and Engineering Research Council (Canada) and the University of Edinburgh School of Biological Sciences. B.C. was supported by the Royal Society (United Kingdom).

#### LITERATURE CITED

- ADAMS, M. D., S. E. CELNIKER, R. A. HOLT, C. A. EVANS, J. D. GOCAYNE *et al.*, 2000 The genome sequence of *Drosophila melanogaster*. *Science* **287**: 2185–2195.
- ASHBURNER, M., K. G. GOLIC and R. S. HAWLEY, 2004 *Drosophila: A Laboratory Handbook*, Ed. 2. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- BACHTROG, D., 2003 Accumulation of *Spock* and *Worf*, two novel non-LTR retrotransposons, on the neo-Y chromosome of *Drosophila miranda*. *Mol. Biol. Evol.* **20**: 173–181.
- BARTOLOMÉ, C., and X. MASIDE, 2004 The lack of recombination drives the fixation of transposable elements on the fourth chromosome of *Drosophila melanogaster*. *Genet. Res.* **83**: 91–100.
- BARTOLOMÉ, C., X. MASIDE and B. CHARLESWORTH, 2002 On the abundance and distribution of transposable elements in the genome of *Drosophila melanogaster*. *Mol. Biol. Evol.* **19**: 926–937.
- BERGMAN, C. M., and D. BENSASSON, 2007 Recent LTR retrotransposon insertion contrasts with waves of non-LTR insertion since speciation in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* **104**: 11340–11345.
- BERGMAN, C. M., H. QUESNEVILLE, D. ANXOLABÉHÈRE and M. ASHBURNER, 2006 Recurrent insertion and duplication generate networks of transposable element sequences in the *Drosophila melanogaster* genome. *Genome Biol.* **7**: R112.
- BIÉMONT, C., A. TSITRONE, C. VIEIRA and C. HOOGLAND, 1997 Transposable element distributions in *Drosophila*. *Genetics* **147**: 1997–1999.
- BROOKFIELD, J. F. Y., 1991 Models of repression of transposition in *P-M* hybrid dysgenesis by *P* cytotype and by zygotically encoded repressor proteins. *Genetics* **128**: 471–486.
- BROOKFIELD, J. F. Y., and R. M. BADGE, 1997 Population genetic models of transposable elements. *Genetica* **100**: 281–294.
- BUTCHER, D., 1995 Muller's ratchet, epistasis, and mutation effects. *Genetics* **141**: 431–437.
- CARR, M., J. R. SOLOWAY, T. E. ROBINSON and J. F. Y. BROOKFIELD, 2001 An investigation of the cause of low variability on the fourth chromosome of *Drosophila melanogaster*. *Mol. Biol. Evol.* **18**: 2260–2269.
- CASALS, F., J. GONZÁLEZ and A. RUIZ, 2006 Abundance and chromosomal distribution of six *Drosophila buzzatii* transposons: *BuT1*, *BuT2*, *BuT3*, *BuT4*, *BuT5*, and *BuT6*. *Chromosoma* **115**: 403–412.
- CHARLESWORTH, B., 1990 Mutation-selection balance and the evolutionary advantage of sex and recombination. *Genet. Res.* **55**: 199–221.
- CHARLESWORTH, B., 1996 Background selection and patterns of genetic diversity in *Drosophila melanogaster*. *Genet. Res.* **68**: 131–149.
- CHARLESWORTH, B., and D. CHARLESWORTH, 1983 The population dynamics of transposable elements. *Genet. Res.* **42**: 1–27.
- CHARLESWORTH, B., and C. H. LANGLEY, 1991 Population genetics of transposable elements in *Drosophila*, pp. 150–176 in *Evolution at the Molecular Level*, edited by R. K. SELANDER, A. G. CLARK and T. S. WHITTAM. Sinauer Associates, Sunderland, MA.
- CHARLESWORTH, B., A. LAPID and D. CANADA, 1992 The distribution of transposable elements within and between chromosomes in a population of *Drosophila melanogaster*. II. Inferences on the nature of selection against elements. *Genet. Res.* **60**: 115–130.
- CHARLESWORTH, B., P. SNEGOWSKI and W. STEPHAN, 1994 The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* **371**: 215–220.
- CHARLESWORTH, B., C. H. LANGLEY and P. D. SNEGOWSKI, 1997 Transposable element distributions in *Drosophila*. *Genetics* **147**: 1993–1995.
- COMBADÃO, J., P. R. A. CAMPOS, F. DIONISIO and I. GORDO, 2007 Small-world networks decrease the speed of Muller's ratchet. *Genet. Res.* **89**: 7–18.
- COMERON, J. M., and M. KREITMAN, 2002 Population, evolutionary and genomic consequences of interference selection. *Genetics* **161**: 389–410.
- CRAIG, N. L., R. CRAIGIE, M. GELLERT and A. M. LAMBOWITZ (Editors), 2002 *Mobile DNA II*. ASM, Washington, DC.
- DOLGIN, E. S., and B. CHARLESWORTH, 2006 The fate of transposable elements in asexual populations. *Genetics* **174**: 817–827.
- DOOLITTLE, W. F., and C. SAPIENZA, 1980 Selfish genes, the phenotype paradigm and genome evolution. *Nature* **284**: 601–603.
- DURET, L., G. MARAIS and C. BIÉMONT, 2000 Transposons but not retrotransposons are located preferentially in regions of high recombination rate in *Caenorhabditis elegans*. *Genetics* **156**: 1661–1669.
- EANES, W. F., C. WESLEY and B. CHARLESWORTH, 1992 Accumulation of *P* elements in minority inversions in natural populations of *Drosophila melanogaster*. *Genet. Res.* **59**: 1–9.
- EWENS, W. J., 2004 *Mathematical Population Genetics*, Ed. 2. Springer-Verlag, Berlin.
- FELSENSTEIN, J., 1974 The evolutionary advantage of recombination. *Genetics* **78**: 737–756.
- FINNEGAN, D. J., 1992 Transposable elements, pp. 1096–1107 in *The Genome of Drosophila melanogaster*, edited by D. L. LINDSLEY and G. ZIMM. Academic Press, New York.
- FONTANILLAS, P., D. L. HARTL and M. REUTER, 2007 Genome organization and gene expression shape the transposable element

- distribution in the *Drosophila melanogaster* euchromatin. *PLoS Genet.* **3**: e210.
- FULLERTON, S. M., A. BERNARDO CARVALHO and A. G. CLARK, 2001 Local rates of recombination are positively correlated with GC content in the human genome. *Mol. Biol. Evol.* **18**: 1139–1142.
- GOLDMAN, A. S., and M. LICHTEN, 1996 The efficiency of meiotic recombination between dispersed sequences in *Saccharomyces cerevisiae* depends upon their chromosomal location. *Genetics* **144**: 43–55.
- GOLDMAN, A. S., and M. LICHTEN, 2000 Restriction of ectopic recombination by interhomolog interactions during *Saccharomyces cerevisiae* meiosis. *Proc. Natl. Acad. Sci. USA* **97**: 9537–9542.
- GORDO, I., and B. CHARLESWORTH, 2000 The degeneration of asexual haploid populations and the speed of Muller's ratchet. *Genetics* **154**: 1379–1387.
- GORDO, I., and B. CHARLESWORTH, 2001 The speed of Muller's ratchet with background selection and the degeneration of Y chromosomes. *Genet. Res.* **78**: 149–161.
- GORDO, I., A. NAVARRO and B. CHARLESWORTH, 2002 Muller's ratchet and the pattern of variation at a neutral locus. *Genetics* **161**: 835–848.
- HILL, W. G., and A. ROBERTSON, 1966 The effect of linkage on limits to artificial selection. *Genet. Res.* **8**: 269–294.
- HOOGLAND, C., and C. BIÉMONT, 1996 Chromosomal distribution of transposable elements in *Drosophila melanogaster*: test of the ectopic recombination model for maintenance of insertion site number. *Genetics* **144**: 197–204.
- HUA-VAN, A., A. LE ROUZIC, C. MAISONHAUTE and P. CAPY, 2005 Abundance, distribution and dynamics of retrotransposable elements and transposons: similarities and differences. *Cytogenet. Genome Res.* **110**: 426–440.
- KAMINKER, J. S., C. M. BERGMAN, B. KRONMILLER, J. CARLSON, R. SVIRSKAS *et al.*, 2002 The transposable elements of the *Drosophila melanogaster* euchromatin: a genomics perspective. *Genome Biol.* **3**: RESEARCH0084.
- KEIGHTLEY, P. D., and S. P. OTTO, 2006 Interference among deleterious mutations favours sex and recombination in finite populations. *Nature* **443**: 89–92.
- KIDWELL, M. G., 2002 Transposable elements and the evolution of genome size in eukaryotes. *Genetica* **115**: 49–63.
- KONDRASHOV, A. S., 1994 Muller's ratchet under epistatic selection. *Genetics* **136**: 1469–1473.
- KREITMAN, M., 1983 Nucleotide polymorphism at the alcohol dehydrogenase locus of *Drosophila melanogaster*. *Nature* **304**: 412–417.
- LANGLEY, C. H., E. A. MONTGOMERY, R. HUDSON, N. KAPLAN and B. CHARLESWORTH, 1988 On the role of unequal exchange in the containment of transposable element copy number. *Genet. Res.* **52**: 223–235.
- LE ROUZIC, A., and G. DECÉLIERE, 2005 Models of the population genetics of transposable elements. *Genet. Res.* **85**: 171–181.
- LE ROUZIC, A., T. S. BOUTIN and P. CAPY, 2007 Long-term evolution of transposable elements. *Proc. Natl. Acad. Sci. USA* **104**: 19375–19380.
- LYNCH, M., 2007 *The Origins of Genome Architecture*. Sinauer Associates, Sunderland, MA.
- MASIDE, X., S. ASSIMACOPOULOS and B. CHARLESWORTH, 2000 Rates of movement of transposable elements on the second chromosome of *Drosophila melanogaster*. *Genet. Res.* **75**: 275–284.
- MASIDE, X., C. BARTOLOMÉ, S. ASSIMACOPOULOS and B. CHARLESWORTH, 2001 Rates and movement of transposable elements on the second chromosome of *Drosophila melanogaster*: in situ hybridization versus Southern blotting data. *Genet. Res.* **78**: 121–136.
- MASIDE, X., S. ASSIMACOPOULOS and B. CHARLESWORTH, 2005 Fixation of transposable elements in the *Drosophila melanogaster* genome. *Genet. Res.* **85**: 195–203.
- MATSUMOTO, M., and T. NISHIMURA, 1998 Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Trans. Mod. Comput. Simul.* **8**: 3–30.
- MCVEAN, G. A. T., and B. CHARLESWORTH, 2000 The effects of Hill–Robertson interference between weakly selected mutations on patterns of molecular evolution and variation. *Genetics* **155**: 929–944.
- MONTGOMERY, E. A., B. CHARLESWORTH and C. H. LANGLEY, 1987 A test for the role of natural selection in the stabilization of transposable element copy number in a population of *Drosophila melanogaster*. *Genet. Res.* **49**: 31–41.
- MONTGOMERY, E. A., S. M. HUANG, C. H. LANGLEY and B. H. JUDD, 1991 Chromosome rearrangement by ectopic recombination in *Drosophila melanogaster*: genome structure and evolution. *Genetics* **129**: 1085–1098.
- MULLER, H. J., 1964 The relation of recombination to mutational advance. *Mutat. Res.* **1**: 2–9.
- NUZHDIIN, S. V., 1999 Sure facts, speculations, and open questions about evolution of transposable elements. *Genetica* **107**: 129–137.
- NUZHDIIN, S. V., and T. F. C. MACKAY, 1995 The genomic rate of transposable element movement in *Drosophila melanogaster*. *Mol. Biol. Evol.* **12**: 180–181.
- ORGEL, L. E., and F. H. CRICK, 1980 Selfish DNA: the ultimate parasite. *Nature* **284**: 604–607.
- PASYUKOVA, E. G., S. V. NUZHDIIN and D. A. FILATOV, 1998 The relationship between the rate of transposition and transposable element copy number for *copia* and *Doc* retrotransposons of *Drosophila melanogaster*. *Genet. Res.* **72**: 1–11.
- RIZZON, C., G. MARAIS, M. GOUY and C. BIÉMONT, 2002 Recombination rate and the distribution of transposable elements in the *Drosophila melanogaster* genome. *Genome Res.* **12**: 400–407.
- SÁNCHEZ-GARCÍA, A., X. MASIDE and B. CHARLESWORTH, 2005 High rate of horizontal transfer of transposable elements in *Drosophila*. *Trends Genet.* **21**: 200–203.
- SHELD AHL, L. A., D. M. WEINREICH and D. M. RAND, 2003 Recombination, dominance and selection on amino acid polymorphism in the *Drosophila* genome: contrasting patterns on the X and fourth chromosomes. *Genetics* **165**: 1195–1208.
- SNIEGOWSKI, P., and B. CHARLESWORTH, 1994 Transposable element numbers in cosmopolitan inversions from a natural population of *Drosophila melanogaster*. *Genetics* **137**: 815–827.
- SÖDERBERG, R. J., and O. G. BERG, 2007 Mutational interference and the progression of Muller's ratchet when mutations have a broad range of deleterious effects. *Genetics* **177**: 971–986.
- STEINEMANN, M., and S. STEINEMANN, 1998 Enigma of Y chromosome degeneration: neo-Y and neo-X chromosomes of *Drosophila miranda* a model for sex chromosome evolution. *Genetica* **102/103**: 409–420.
- STEPHAN, W., and Y. KIM, 2002 Recent applications of diffusion theory to population genetics, pp. 72–93 in *Modern Developments in Theoretical Population Genetics*, edited by M. SLATKIN and M. VEUILLE. Oxford University Press, Oxford.
- SUH, D. S., E. H. CHOI, T. YAMAZAKI and K. HARADA, 1995 Studies on the transposition rates of mobile genetic elements in a natural population of *Drosophila melanogaster*. *Mol. Biol. Evol.* **12**: 748–758.
- VIEIRA, C., and C. BIÉMONT, 1997 Transposition rate of the 412 retrotransposable element is independent of copy number in natural populations of *Drosophila simulans*. *Mol. Biol. Evol.* **14**: 185–188.
- WRIGHT, S. I., and D. J. SCHOEN, 1999 Transposon dynamics and the breeding system. *Genetica* **107**: 139–148.